

## RESEARCH INTEREST

I'm interested in language and speech processing. In particular, I am enthusiastic about building efficient and robust speech recognition systems in low-resource and multilingual settings. I am also interested in understanding the intricacies of human-machine interactions towards building intelligent multimodal dialogue systems. Presently, I am focussed on building speech and language data resources that would enable research and development of language technology systems for Zambian languages.

---

## EDUCATION

- MSc., Mathematical Sciences**, Africa Institute of Mathematical Sciences, Kigali, Rwanda Oct 2019 – **Aug 2020**
- Supervisor: Prof. Antonios Anastasopoulos, George Mason University, U.S.A
  - Specialization: **Machine Intelligence**
- MSc., Information Technology Management**, Binary University, Puchong, Malaysia Apr 2013 – Oct 2014
- Specialization: **Information Technology**
- BSc., Computer Science**, University of Zambia, Lusaka, Zambia Apr 2007 – Aug 2011
- Specialization: **Computer Systems Engineering**
- 

## TEACHING EXPERIENCE

- Lecturer**, Department of Computer Science, University of Zambia, Zambia May 2015 – **Present**
- **Bachelor of Computer Science courses:**

CSC 2011 - Computer Systems	2021/2022
CSC 2702 - Databases & Information System	2021/2022
CSC 3712 - Advanced Databases	2021/2022
CSC 4792 - Data Mining and Warehousing	2021/2022
CSC 4000 - Final Year Project Supervision	2021/2022
  - **Postgraduate Diploma in Computer Science courses:**

CSC 5711 - Databases & Information System	2021/2022
---	-----------
- Instructor**, CISCO Academy, University of Zambia, Zambia Sep 2016 – Jun 2019
- Instructor of Python Programming,
  - IT Essentials and Cisco Certified Network Associate (CCNA) Level 1 and Level 2 CISCO courses.
- Teaching Assistant**, Department of Computer Science, University of Zambia, Zambia Feb 2012 – May 2015
- Teaching assistant (Tutor) for Database Systems courses under the Staff Development Fellowship (SDF).
- 

## PERSONAL DEVELOPMENT

- **Mathematics for Machine Learning Specialization** by Imperial College London on Coursera Nov 2019  
Certificate of completion
- **Deep Learning Specialization** by Deeplearningai on Coursera Aug 2018  
Certificate of completion
- **Machine Learning** by Stanford University on Coursera Jul 2018  
Certificate of completion

---

## SELECTED PUBLICATIONS

- **C. Sikasote**, K. Siaminwe, S. Mwape, B. Zulu, M. Phiri, M.C. Phiri, D. Zulu, M. Nyirenda, A. Anastasopoulos, *Zambezi Voice: A Multilingual Speech Corpus for Zambian Languages*, INTERSPEECH, 2023.
- **C. Sikasote**, E. Mukonde, M.M Alam, A. Anastasopoulos, *BIG-C: a Multimodal Multi-Purpose Dataset for Bemba*, Association for Computational Linguistics (ACL), 2023.
- **C. Sikasote**, A. Anastasopoulos, *BembaSpeech: A Speech Recognition Corpus for the Bemba Language*, Proceedings of the Language and Resource Evaluation Conference (LREC), 2022.
- J. Kreutzer, I. Caswell, L. Wang, A. Wahab, D.V Esch, N. Ulzii-Orshikh, A. Tapo, N. Subramani, A. Sokolov, **C. Sikasote**, M. Setyawan, S. Sarin, S. Samb, B. Sagot, C. Rivera, A. Rios, I. Papadimitriou, S. Osei, P.O Suarez, I. Orife, K. Ogueji, A. Rubungo, T. Q. Nguyen, M. Müller, A. Müller, S.H. Muhammad, N. Muhammad, A. Mnyakeni, J. Mirzakhlov, T. Matangira, C. Leong, N. Lawson, S. Kudugunta, Y. Jernite, M. Jenny, O. Firat, B.F.P Dossou, S. Dlamini, N. Silva, S.Bali, S. Biderman, A. Battisti, A. Baruwa, A. Bapna, P. Baljekar, I. A. Azime, A. Awokoya, D. Ataman, O. Ahia, S. Agrawal, M. Adeyemi, *Quality at a glance: An audit of web-crawled multilingual datasets*, Transactions of the Association for Computational Linguistics (TACL), 2022
- S. H. Chiwamba, J. Phiri, P. O.Y. Nkunika, **C. Sikasote**, M. M. Kabemba, M. N. Moonga, *Automated Fall Armyworm (Spodoptera frugiperda, J.E. Smith) Pheromone Trap Based on Machine Learning*, Journal of Computer Science 2019, 15 (12): 1759.1779

---

## SCHOLARSHIPS AND FELLOWSHIPS

- **African Master of Machine Intelligence Scholarship**, Kigali, Rwanda Sep 2019
- **Special Research Fellowship**, University of Zambia, Lusaka, Zambia Sep 2019
- **Binary University Scholarship**, Binary University, Malaysia Apr 2013
- **Staff Development Fellowship**, University of Zambia, Lusaka, Zambia Feb 2012
- **Zambia Higher Education Scholarship**, University of Zambia, Lusaka, Zambia Apr 2007

---

## GRANTS AND AWARDS

- **Recipient**, Diversity and Inclusion Virtual Award, ACL 2023 Jun 2023
- **Recipient**, ISCA-Travel Grant, INTERSPEECH 2023 Jun 2023
- **Co-PI**, Multimodal Datasets for the Bemba Language Project Apr 2021 - Aug 2022

---

## TECHNICAL SKILLS

- **Programming Languages:** Python
- **Database technologies:** MS SQL, MySQL
- **Operating System platforms:** Windows, Linux
- **Machine Learning libraries:** Scikit-learn, Pandas, NumPy, SciPy
- **Data Visualization tools:** Matplotlib, tSNE
- **Deep Learning frameworks:** Pytorch, Tensorflow
- **Speech Recognition Tools:** DeepSpeech, FairSeq

---

## OPEN-SOURCE DATASETS

- **Zambezi Voice** : [Lead researcher and creator] of a multilingual speech corpus containing speech recognition data resources for native African languages of Zambia. [ON-GOING]  
Link: <https://github.com/unza-speech-lab/zambezi-voice>
- **Multimodal Dataset for the Bemba** : [Lead researcher and Co-PI] of a large scale multimodal dataset for Bemba language containing multi-turn image-grounded dialogues amounting to over 180 hours of speech with corresponding Bemba transcriptions and translation in English. The project was funded by *Lacuna Fund*.  
Link: <https://github.com/csikasote/bigc>
- **BembaSpeech** : [Lead researcher and creator] an ASR corpus consisting of 24 hrs of read speech for the Bemba language spoken by over 30% of the population in Zambia.  
Link: <https://github.com/csikasote/BembaSpeech>
- **Multilingual ASR Corpus for African Languages** : [Contributor] to a multilingual speech corpus consisting of over 60 hours of read speech for 19 African languages. I contributed 2 hours of read speech for Bemba language.  
Link: <https://github.com/besacier/AMMIcourse/tree/master/STUDENTS-RETURN>

---

## STUDENT SUPERVISION AND MENTORING

- |  |           |
|--|-----------|
| • <b>Nawiti Simwanza</b> , Undergraduate, University of Zambia           | 2022/2023 |
| • <b>Felix Kunda</b> , Undergraduate, University of Zambia               | 2022/2023 |
| • <b>Ngonidzashe Mbita Sikazwe</b> , Undergraduate, University of Zambia | 2022/2023 |
| • <b>Stanly Mwape</b> , Undergraduate, University of Zambia              | 2021/2022 |
| • <b>Kalinda Siaminwe</b> , Undergraduate, University of Zambia,         | 2021/2022 |

---

## COMMUNITY SERVICE

- |  |          |
|--|----------|
| • <b>Conference Reviewer</b>   |          |
| Pan African Conference on Science, Computing & Telecommunications (PACT)     | Sep 2023 |
| • <b>Conference Volunteer</b>  |          |
| International Conference on Learning Representations (ICLR)                  | Mar 2021 |
| International Conference on Artificial Intelligence and Statistics (AISTATS) | Apr 2021 |